James Z. Wang
Reginald B. Adams, Jr. *Editors*

Modeling Visual Aesthetics, Emotion, and Artistic Style

Springer

Modeling Visual Aesthetics, Emotion, and Artistic Style

James Z. Wang • Reginald B. Adams, Jr.
Editors

Modeling Visual Aesthetics, Emotion, and Artistic Style

Springer

Editors
James Z. Wang
Information Sciences and Technology
The Pennsylvania State University
University Park, PA, USA

Reginald B. Adams, Jr.
Department of Psychology
The Pennsylvania State University
University Park, PA, USA

ISBN 978-3-031-50268-2 ISBN 978-3-031-50269-9 (eBook)
https://doi.org/10.1007/978-3-031-50269-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

We dedicate this book to our families, whose love and support have been our constant source of strength and inspiration throughout the entire editing process.

James and Reg

Foreword

Over millennia, people have learned both overt behaviors and more subtle cues to communicate to one another. Anthropological motivations abound, from the need to survive, to protect one's community, and to secure and exploit resources. Along the evolutionary path, people had to interact with others. Even before the emergence of languages, humans had to sense, read, and assign meaning, mostly instantaneously, to observed and perceived facial expressions, body gestures, and actions. Whatever physiological processes were responsible for these abilities, they are now embedded in our genetic makeup and allow children to acquire, learn, and ascribe meaning to social environments.

Over the past century or so, first psychologists, and then computer scientists, began quantitative studies and experiments to try to understand how socially communicative behaviors arose. Fundamental questions about human perception and how sensory systems might be "wired" dominated pre-computational studies. When computers, visual sensors, and displays emerged to support scientific endeavors, new foci arose on linking synthetic perception, algorithmic models of social phenomena, and generalized mechano-robotic and graphical generation of social signals. In other words, the era of synthesized virtual humans began.

I was extremely fortunate to be able to participate in and contribute to this new field from my PhD thesis at the University of Toronto in 1974 up to the present. By 1990, there were several robust research communities interested in human perception, robotics, computer vision, and computer graphics. David Zeltzer of MIT, Brian Barsky of the University of California at Berkeley, and myself from the University of Pennsylvania, organized a "Workshop on the Mechanics, Control and Animation of Articulated Figures" held at the MIT Media Lab in April 1989. We invited participants from multiple perspectives to share their thoughts on humans and their virtual or robotic embodiments. This successful cross-disciplinary meeting led to the first book in 1990—*Making Them Move: Mechanics, Control and Animation of Articulated Figures*—to encompass these disparate but ultimately deeply connected viewpoints.

This volume on *Modeling Visual Aesthetics, Emotion, and Artistic Style*, thoughtfully curated by James Z. Wang and Reginald B. Adams, Jr., is a perfect bookend to that earlier *Making Them Move* collection. Human behavior observation has been dramatically enabled by low-cost, high-resolution image acquisition hardware feeding real-time computer vision motion analysis systems. Computer virtual human simulation and computer graphics have evolved to the point where non-real-time movie actors, characters, and monsters are produced with efficiency and regularity by the movie and game industry, while real-time human agents are now taking on public roles as announcers, influencers, and assistants. Their ubiquity has fostered artistic interest and study of the human aesthetic. Driving the range of contemporary applications are new tools from the Artificial Intelligence and Machine Learning research communities. We could perhaps only dream about these in 1990. They are the new foundation for human perception and simulation research.

With modern computational tools and decades of computer graphics simulations to build on, additional fascinating aspects of human communication can be studied, modeled, and reproduced. Human emotions, and their companion attributes of mood and personality expressed by face and gesture, have long been of interest to multiple research communities, including the social sciences as well as the computational ones. This volume addresses emotional displays and understanding, including novel dimensions such as threats, which are of clear evolutionary value. As its title aptly describes, this volume also includes new considerations of aesthetics and artistic style. Critical questions of bias and sexual discrimination must be addressed as learning systems depend on datasets that might, inadvertently or naively, perpetuate stereotypes, cultural misconceptions, or prejudices. The maturity of the underlying computational foundations now admits these humanistic questions. A number of works in this volume explore this space of unique human characteristics.

James and Reg have assembled an outstanding collection of current approaches to modeling novel human dimensions. It will be a classic of interdisciplinary computational studies. Enjoy!

Haverford, PA, USA Norman I. Badler
June 2023

Preface

Visual aesthetics, emotional expression, and artistic style are essential components of human perception and experience, and their significance has only grown with the increasing prevalence of digital media and technologies. The ability to computationally model and analyze these complex concepts has been a longstanding goal in the fields of computer vision, affective computing, and robotics. This timely book represents the collective efforts of active researchers from a diverse set of fields, including computer vision, robotics, psychology, graphics, data mining, machine learning, movement analysis, and art history, who have come together to address these challenging and critical research questions. As our world becomes more interconnected and reliant on digital platforms and artificial intelligence, understanding and effectively utilizing these aspects of human experience has become increasingly important, making this book a vital resource for both researchers and practitioners alike.

The chapters of this book cover a wide range of topics related to the computational modeling of aesthetics, emotion, and artistic style. The first part provides background knowledge related to emotion models and machine learning. The next two parts explore social visual perception in humans and its application to computer vision. Specifically, Part II lays the groundwork by discussing the basic psychological and neurological underpinnings of social and emotional perception from faces and bodies. Part III extends this understanding into the realm of technology, demonstrating methods to train computer systems to detect discrete and micro-momentary emotional expressions from facial and body cues, question the notion of facial neutrality, and broaden the scope of research to include children as well as adults in the context of emotion perception. Part IV focuses on the dynamic intersection of art and technology, shedding light on the language of photography, the interplay between breath-driven robotic performances and human dance, and the application of machine learning in the contextual analysis of artistic style. The remaining three parts dive deeper into the computational modeling of visual aesthetics, emotion, and artistic style.

One of the unique features of this book is its multidisciplinary approach, bringing together contributions from various domains, such as computer science, psychology, art history, and cognitive science. This interdisciplinary approach fosters a more holistic understanding of the subject matter and encourages cross-disciplinary collaboration, leading to novel insights and advancements in the field.

The versatile nature of the book format has enabled us to encompass an array of contribution types. These include comprehensive tutorials and reviews of theoretical frameworks and computational methodologies, extensive literature surveys, novel methodological approaches, in-depth case studies, insightful opinion pieces, rigorous empirical investigations, and comparative analyses.

Another feature of this book is its focus on cutting-edge research. The methods and information presented in this book represent the latest developments in the field and have the potential to significantly advance the field. The comprehensive and in-depth treatment of topics offered by book chapters provides a richer understanding of the subject for readers, which can be especially beneficial for those new to a new interdisciplinary field or those looking to expand their knowledge.

Finally, the impact of the results presented in the book can be far-reaching. The ability to computationally model aesthetics, emotion, and artistic style has the potential to enable many computer and robotic applications that can benefit millions of people around the world. From children needing care to the elderly needing assistance, from amateur photographers to people working alongside robots, the impact of this work is broad.

We trust that this book will serve as a valuable resource for researchers, practitioners, educators, and students who are interested in advancing the field. The cross-disciplinary nature of the book increases the chances of a wider audience accessing the research, leading to broader dissemination and long-term recognition of the presented findings. We hope that this book will inspire further research, foster interdisciplinary collaboration, and contribute to the advancement of computational modeling of visual aesthetics, emotion, and artistic style.

State College, PA, USA — James Z. Wang
June 2023 — Reginald B. Adams, Jr.

Acknowledgments

First and foremost, we would like to express our deepest gratitude to our colleagues who contributed their invaluable expertise, knowledge, and insights to this book. Their dedication and commitment to advancing this field have substantially enriched the content and elevated the overall quality of this work. It has been a privilege to collaborate with such an accomplished and diverse assembly of researchers in this endeavor.

We convey our sincere appreciation to the editorial team at Springer Nature for their steadfast support, professionalism, and guidance throughout the publication process. Their invaluable feedback and constructive suggestions have played an important role in shaping the final product. Special thanks go to our editor, Susan E. Grove, and project coordinator, Arun S. Shanmugam, for their relentless enthusiasm and encouragement, which motivated us to strive for excellence in our work. Additionally, we would like to express our gratitude to the anonymous reviewers for their valuable insights and constructive feedback on our book proposal.

We deeply appreciate Norman Badler for his insightful Foreword. His rich experiences have helped contextualize the evolution of our field, emphasizing the roles of artificial intelligence and machine learning in human perception and simulation research. His emphasis on addressing issues like bias and cultural misconceptions in learning systems is invaluable.

We would also like to thank our academic mentors, advisees, colleagues, and collaborators who have inspired and supported our research over the years. Their expertise, encouragement, and friendship have been essential in the development of our understanding and passion for the field. In particular, J. Z. Wang is grateful to Gio Wiederhold, Dennis A. Hejhal, Martin A. Fischler, and Edward H. Shortliffe for their invaluable guidance, wisdom, and belief in his potential. He is also grateful for the support and encouragement received from Adam Fineberg, Yelin Kim, Tatiana D. Korelsky, and Juan P. Wachs. R. B. Adams, Jr. is particularly grateful to Robert E. Kleck, Ursula Hess, and Nalini Ambady for their early mentorship and encouragement. He is also grateful to all his colleagues in psychology and vision science who have helped him champion the field of Social Vision. Finally, he is grateful to J. Z. Wang for spearheading this book, for collaboration over the years extending his own work in new directions, and enabling him to apply insights from social visual perception to this burgeoning field of computer vision.

Our deepest appreciation goes to the countless artists and creators whose work has inspired and fueled our research. Their artistic expressions and creative pursuits provide the foundation for our exploration of this fascinating field.

We gratefully acknowledge the financial support provided by the National Science Foundation (NSF) under Grant Nos. 1110970, 1921783, and 2234195, which has been instrumental in advancing the research presented in this book. This funding has enabled us to pursue innovative research directions, collaborate with leading experts, and ultimately contribute to the growing body of knowledge in this field. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. Additionally, J. Z. Wang extends his sincere appreciation to the Amazon Research Awards program for their gift, which has further bolstered the research of his team and facilitated the development of cutting-edge methodologies and technologies. His team's research in visual art has been supported in part by the National Endowment for the Humanities (NEH) under Grant Nos. HAA-271801-20 and HAA-287938-22. His team's research in machine learning has been supported in part by the NSF under Grant Nos. 2205004 and 2216127. We are immensely grateful for the confidence and investment these organizations have placed in our work, and we strive to continue making meaningful contributions to this exciting and dynamic field.

Last but not least, we would like to express our heartfelt gratitude to our family and friends for their constant love, support, and encouragement. Their belief in our abilities, patience with our countless hours of work, and persistent understanding have been the bedrock of our endeavors. Without them, this book would not have been possible. J. Z. Wang would like to especially thank Jia Li and their children Justina and Nora Wang for their inspiration and consistent understanding. Similarly, R. B. Adams, Jr. thanks Katharine Donnelly Adams as well as their children Henry and Lena Logan Adams for their understanding and encouragement throughout the process of putting this book together.

Contents

Contributors

Reginald B. Adams, Jr. is a Professor of Psychology at The Pennsylvania State University. He received his Ph.D. in Social Psychology from Dartmouth College in 2002. Reg is interested in how we extract social and emotional meaning from nonverbal cues, particularly via the face. His work addresses how multiple social messages (e.g., emotion, gender, race, age, etc.) combine across multiple modalities and interact to form the unified representations that guide our impressions of and responses to others. Although his questions are social psychological in origin, his research draws upon vision cognition and affective neuroscience to address social perception at the functional and neuroanatomical levels. With his colleagues, Reg helped establish and champion the subfield of Social Vision by publishing an edited volume titled *The Science of Social Vision* (Adams, Ambady, Nakayama, & Shimojo, 2010, Oxford University Press). His research has been funded by NSF, NIA, and NIMH (NIH).

Daniel N. Albohn is a Principal Researcher at the University of Chicago Booth School of Business. Dan received his Ph.D. in Psychology from The Pennsylvania State University. His research uses data-driven, machine learning, and human responses to examine how social cues inform judgments of people and objects. He has a particular interest in how individuals extract information from neutral or minimally expressive faces.

Norman I. Badler is an Emeritus Professor of Computer and Information Science at the University of Pennsylvania. He received his B.A. in Creative Studies Mathematics from the University of California Santa Barbara in 1970, his M.S. in Mathematics from the University of Toronto in 1971, and his Ph.D. in Computer Science from the University of Toronto in 1975. His research has involved developing software to acquire, simulate, animate, and control 3D computer graphics human body, face, gesture, locomotion, and manual task motions, both individually and for heterogeneous groups. These virtual humans are meant to portray physical, cognitive, perceptual, personality, relationship, and cultural parameters. He has supervised or co-supervised 61 Ph.D. students, many of whom have become academics or researchers in the movie visual effects and game industries. He was the founding Director of the SIG Center for Computer Graphics, the Center for Human Modeling and Simulation, and the ViDi Center for Digital Visualization at Penn. He has co-authored five books, one in digital human modeling and the other four in virtual crowd simulation. He serves part-time as the Head of Metaverse Research at Philadelphia-based Cesium GS Inc.

Joseph C. Brandenburg is a fourth-year graduate student in the school psychology program at The Pennsylvania State University. He has been working in and collaborating with the members of the Social Vision and Interpersonal Perception lab for 9+ years throughout his undergraduate and postgraduate career. Joe has his Master of Science in Clinical Psychology from Millersville University and his Master of Education in School Psychology. Joe's research interests include emotion perception, emotion regulation, stress, and psychophysiology. He has worked on myriad projects including these topics with an emerging interest in how wearable technologies can help within these already existing arenas of interest.

Kathryn Brown is an Associate Professor of Art History at Loughborough University (UK). Her books include *Women Readers in French Painting 1870–1890* (2012), *Matisse's Poets: Critical Performance in the Artist's Book* (2017), *Henri Matisse* (2021), and *Dialogues with Degas: Influence and Antagonism in Contemporary Art* (2023). She has edited several essay collections, including *Digital Humanities and Art History* (Routledge, 2020). Brown's research has been supported by numerous funders including the Association of Art History (UK), the British Academy, the Independent Social Research Foundation, and the Terra Foundation for American Art. In 2021, Brown was a Paul Mellon Visiting Senior Fellow at the Center for Advanced Study in the Visual Arts (Washington, DC).

Pamala N. Dayley is a third-year Social Psychology graduate student at UCLA, supervised by Dr. Kerri Johnson. She received her Bachelor's degree from The Pennsylvania State University, Abington campus, and her Master's degree from UCLA. Her research interests include the perceptions of others (face and body), judgments made about targets based on perceptual cues, and the downstream consequences (e.g., discrimination) of said judgments. She is a National Science Foundation awardee and a National Defense Science and Engineering Graduate Fellow.

Funda Durupinar received a B.S. degree from Middle East Technical University, Ankara, Turkey and an M.S. degree from Bilkent University in Computer Engineering in 2002 and 2004, respectively. She received her Ph.D. degree in August 2010 from the Department of Computer Engineering at Bilkent University, Ankara, Turkey. After completing her Ph.D., she worked as a Postdoctoral Researcher at the Center for Human Modeling and Simulation, University of Pennsylvania. She worked as a Software Engineer at Memorial Sloan Kettering Cancer Center and as a Senior Research Associate at Oregon Health & Science University. She is currently an Assistant Professor in the Department of Computer Science at the University of Massachusetts at Boston. Her research links computer graphics, artificial intelligence, and psychology with a focus on creating believable virtual humans.

Niki Efthymiou is a Ph.D. student at the School of Electrical and Computer Engineering, National Technical University of Athens (NTUA), under the supervision of Prof. Petros Maragos. She is working primarily in computer vision problems associated during Human–Robot Interaction. She is a Researcher at the Computer Vision, Speech Communication, and Signal Processing Group at NTUA, and her research interests lie in the fields of gesture, action, and emotion recognition, with a focus on Child–Robot Interaction. She received her Diploma degree in Applied Mathematics and Master's degree in Computational Mechanics from NTUA.

Panagiotis P. Filntisis is a Postdoctoral Researcher at the IRAL lab of the National Technical University of Athens and a Research Assistant at the Athena Research and Innovation Center. He received his Ph.D. in 2022 under the supervision of Prof. Petros Maragos and holds an M.Eng. Diploma degree in ECE from NTUA. His work lies at the crossroads of computer vision and audio processing for affective computing.

Franz Götz-Hahn is currently working as a Postdoctoral Researcher in the Intelligent Embedded Systems group at the University of Kassel, where he heads the AI for Motion research group. He received his M.Sc. in Artificial Intelligence from Maastricht University, Netherlands and the Ph.D. in Computer Science from the University of Konstanz, Germany with his thesis titled "Video Quality Assessment in-the-wild." Franz's dissertation was the culmination of pioneering work in the field of deep learning for image and video quality assessment, including the (co-)authorship of KonVid-1k and KonVid-150k, two of the most influential and largest in-the-wild video quality datasets to date. Recently, he has expanded his expertise beyond image and video quality toward using artificial intelligence more generally in domains involving motion, such as in automotive.

Uğur Güdükbay received a B.S. degree in Computer Engineering from the Middle East Technical University, Ankara, Turkey, in 1987 and an M.S. and Ph.D. degrees in Computer Engineering and Information Science from Bilkent University, Ankara, Turkey, in 1989 and 1994, respectively. He conducted research as a Postdoctoral Fellow at the Human Modeling and Simulation Laboratory at the University of Pennsylvania. Currently, he is a Professor in the Department of Computer Engineering at Bilkent University. His research interests include human modeling and animation, conversational virtual agents, personality and emotion synthesis, crowd simulation, rendering, and visualization. He is a senior member of ACM and IEEE.

Nicole Hedgecoth is a doctoral candidate at The Pennsylvania State University. They earned their Bachelor's degree in Psychology and their first Master's degree in Negotiation and Conflict management from the University of Baltimore, before continuing to earn their second Master's in Psychology at The Pennsylvania State University. Nikki's research interest is in bringing an interdisciplinary approach to social vision and person perception, drawing on feminist theory and methods to inform their work.

Vlad Hosu is a Postdoctoral Researcher in the Multimedia Signal Processing group at the University of Konstanz. With a Ph.D. in Computer Vision from the National University of Singapore, his dissertation focused on the aesthetics of lighting design in computational photography. Vlad is dedicated to exploring the intersection of technical and aesthetic quality assessment by studying human visual perception. He is developing innovative visual quality models using machine-learning techniques and crowdsourcing. Vlad's contributions include co-authoring several central databases for the field and widely used predictive models.

Yifan Jiang is a Ph.D. student in the Department of Electrical and Computer Engineering at the University of Texas at Austin, supervised by Prof. Zhangyang (Atlas) Wang. He received his Bachelor's degree from Huazhong University of Science and Technology, Wuhan, China. His research interests range from neural rendering, generative models, and computational photography. He also completed internships at Bytedance AI Lab, Adobe, and Google Research. He is a recipient of the 2023 Apple Scholar in AI/ML.

Kerri L. Johnson is a Professor in the departments of Communication and Psychology at UCLA, where she currently serves as the Associate Vice Chancellor for Faculty Development. After receiving her Ph.D. from Cornell University in 2004, she was a Postdoctoral Fellow at NYU before joining the faculty at UCLA. Her research is at the forefront of the burgeoning field of Social Vision. She examines how the perception of cues in the face and body impacts interpersonal judgments, behaviors, and biases. She has published widely on how such perceptions inform a range of downstream judgments, including politics, sexual orientation, and even religion, often documenting profound biases in how appearance impacts meaningful outcomes. Her lab pursues highly interdisciplinary work, including both theoretical and methodological breadth from across the allied social, cognitive, and visual sciences.

Baris Kandemir received a B.Sc. degree in Electrical and Electronics Engineering from Boğaziçi (Boğaziçi) University, Istanbul, Turkey with high honor in 2012. He obtained his Ph.D. degree from the College of Information Sciences and

Technology, The Pennsylvania State University, University Park, in 2019. Since July 2018, he has been with DeepMap, a subsidiary of NVIDIA, where he is a systems software engineer. His main interests are biomedical image processing, computational aesthetics and affect, and 3D computational geometry. During his Ph.D. studies, he investigated visual balance and aesthetics, relationship among visual cues, affect, and demographics. Additionally, he studied the 3D segmentation of tubular structures in confocal microscopy stacks. He is currently working on crowdsourced mapping using NVIDIA's perception modules.

Hanjoo Kim is a Postdoctoral Research Fellow at the Heinz C. Prechter Bipolar Research Program, University of Michigan. He obtained a Ph.D. in Clinical Psychology from The Pennsylvania State University and completed an APA-accredited psychology internship at the New Mexico VA/Southwest Consortium. His primary research interests center around the "underlying mechanisms" of emotional disorders, including anxiety, unipolar depression, and bipolar spectrum disorders. Currently, his research focuses on understanding the emotion dysregulation processes involved in repetitive negative thoughts, such as worry and rumination. To investigate this topic, he is utilizing various psychophysiological methodologies, such as skin conductance, emotional facial expressions, and heart rate variability, alongside intensive longitudinal data analysis.

Kestutis Kveraga is an Assistant Professor at the Harvard Medical School and an Assistant in Neuroscience at the Massachusetts General Hospital. He is a cognitive neuroscientist who studies the neural mechanisms of threat perception from naturalistic stimuli, with strong interests in visual pathway function and autism. He is also interested in neural aesthetics and how brain activity can be employed to predict and shape architectural design and art. He has expertise in neuroimaging methods, such as structural and functional MRI (including ultra-high-field high-resolution 7T fMRI), MEG and EEG, psychophysical techniques (eye and limb tracking, visual pathway biasing), and brain connectivity analyses (e.g., Dynamic Causal Modeling and biomagnetic phase synchrony).

Kate Ladenheim is a choreographer, media designer, and creative technologist who researches bodies in motion and how they impact and are impacted by systems of social and technological pressure. Her work has been presented internationally and spans interactive installations, media design, dance performance, and robotics research. Ladenheim holds an M.F.A. in Media Design Practices from ArtCenter College of Design. She recently assisted robotics research at UCLA, and was the 2019–2020 Artist in Residence at the Robotics, Automation, & Dance (RAD) Lab at UIUC. Her work was celebrated in *Dance Magazine* as one of their "25 to Watch" and "Best of 2018." She is the current Artist in Residence at the Maya Brin Institute for New Performance, a faculty role at the University of Maryland—College Park.

Amy LaViers is the Director of the Robotics, Automation, and Dance (RAD) Lab. Her choreography and machine designs have been presented internationally, including at Joe's Pub at the Public Theater and the Performance Arcade. Her writing has appeared in academic journals like *Nature* and *Robotics and Autonomous Systems* as well as public venues like American Scientist and Aeon. She is a recipient of DARPA's Young Faculty Award (YFA), and her teaching has been recognized on the list of Teachers Ranked as Excellent by Their Students, with outstanding distinction, at the University of Illinois at Urbana-Champaign (UIUC). She has held positions as a co-founder of three start-up companies and as an engineering faculty member at UIUC and the University of Virginia. She holds a CMA from the Laban/Bartenieff Institute of Movement Studies, a Ph.D. and M.S. from Georgia Institute of Technology, and a B.S.E. and certificate in Dance from Princeton University.

Jia Li is a Professor of Statistics and (by courtesy) Computer Science at The Pennsylvania State University. Her research interests include machine learning and image analysis. For her innovations in image retrieval, annotation, aesthetics/composition analysis, and other areas, she has been awarded sixteen US patents. She worked as a Program Director at the National Science Foundation from 2011 to 2013, a Visiting Scientist at Google Labs in Pittsburgh from 2007 to 2008, and a Researcher at the Xerox Palo Alto Research Center from 1999 to 2000. She received an M.Sc. degree in Electrical Engineering (1995), an M.Sc. degree in Statistics (1998), and a Ph.D. degree in Electrical Engineering (1999) from Stanford University. She was Editor-in-Chief of *Statistical Analysis and Data Mining: The ASA Data Science Journal* from 2018 to 2020. She is a Fellow of the Institute of Electrical and Electronics Engineers and a Fellow of the American Statistical Association.

Xin Lu received her Ph.D. degree from the College of Information Sciences and Technology, The Pennsylvania State University, University Park in 2016. Prior to that, she received a B.E. and B.A. degrees in Electronic Engineering and English and an M.E. degree in Signal and Information Processing, all from Tianjin University, China. Since August 2015, she has been with Adobe Inc., where she is currently a Senior Manager and Scientist. Her main research interests are image generation, image segmentation, image aesthetics and emotions, and efficient neural networks. During her Ph.D. studies, she discovered and verified the relationship between simplicity and valence and angularity and valence in complex scenes.

QT Luong is a former computer vision researcher with positions at the University of California, Berkeley, and SRI International turned freelance photographer. His Ph.D. thesis, "Fundamental Matrix and Self-calibration," introduced concepts that spawned a decade of research. The resulting 1992 European Conference on Computer Vision paper "Camera self-calibration: Theory and experiments" (with O. Faugeras and S. Maybank) won the inaugural Koenderink Prize for Fundamental Contributions in Computer Vision in 2008. He is the coauthor (with O. Faugeras) of the book *The Geometry of Multiple Images* (MIT Press 2001/2004). Luong was the first to photograph all of America's 63 national parks—in large format.

He received the Sierra Club's Ansel Adams Award for Photography and the National Parks Conservation Association's Robin W. Winks Award for Enhancing Public Understanding of National Parks. His best-selling book *Treasured Lands: A Photographic Odyssey Through America's National Parks* (2016) won 12 national and international book awards.

Catherine Maguire is a movement educator and dance artist. She is a master teacher of the Laban/Bartenieff Movement System (LBMS) and a Certified Movement Analyst (CMA), having taught and co-coordinated movement analysis certification training programs in the USA, Europe, Mexico, and China. Maguire is a faculty member of WholeMovement, a coterie of movement analysts working together to promote movement studies globally. She has coauthored several publications on expressive robotic systems, including *Making Meaning with Machines: Somatic Strategies, Choreographic Technologies and Notational Abstractions Through a Laban/Bartenieff Lens*. She was the Founder and Artistic Director of Offspring Dance Company in New York City and the Founder and Head of the dance program at Drew University in Madison, NJ, as well as Assistant Professor of dance at Piedmont Virginia Community College. She lives in central Virginia where she teaches ongoing movement classes designed to foster self-expression, body connectivity, and transformation through movement.

Petros Maragos is a full Professor of the School of Electrical and Computer Engineering, National Technical University of Athens, Greece, and Director of the Intelligent Robotics and Automation Lab and the CVSP Group. He has worked as a Professor at USA universities, including Harvard University (1985–93) and Georgia Tech (1993–98). He is also the coordinator of a Robotics Research Unit at the Athena Research and Innovation Center. His research and teaching interests include signal processing and machine learning, computer vision, speech/language, and robotics. He is the recipient of several awards for his academic work. He has been the PI of several US, European, and Greek research projects and served as General Chair of EUSIPCO'17 and ICASSP'23. He is a Fellow of IEEE and EURASIP.

Michelle G. Newman is a Professor of Psychology and Psychiatry and Director of the Center for the Treatment of Anxiety and Depression at The Pennsylvania State University. She received her Ph.D. in Clinical Psychology from the University of Stony Brook in 1992 and completed a postdoctoral fellowship at Stanford University in 1994. Dr. Newman has conducted basic and applied research on anxiety disorders and depression and has published over 200 papers on these topics. She is the past editor of *Behavior Therapy* and is currently Associate Editor of the *Journal of Anxiety Disorders*. She is also the recipient of the APA Division 12 Turner Award for distinguished contribution to clinical research, APA Division 29 Award for Distinguished Publication of Psychotherapy Research Award, ABCT Outstanding Service Award, APA Division 12 Toy Caldwell-Colbert Award for Distinguished

Educator in Clinical Psychology, and Raymond Lombra Award for Distinction in the Social or Life Sciences. She is also a Fellow of the American Psychological Association Divisions 29 and 12, the Association for Behavioral and Cognitive Therapies, and the American Psychological Society.

Flora Oswald is a recent graduate from Penn State's dual-title doctoral program in Psychology and Women's, Gender, and Sexuality Studies, and an Assistant Research Professor at the University of Connecticut. Flora is interested in how marginalized identities shape people's experiences and perceptions of their social worlds, with a particular focus on stereotyping and stigmatization. Much of Flora's current work bridges feminist social psychological approaches with visual perception research to elucidate a feminist social vision perspective that prioritizes marginalized perceivers. Flora's work has been supported by awards from the Social Sciences and Humanities Research Council of Canada, Women and Gender Equality Canada, the Government of Alberta, and the Society for Personality and Social Psychology, among others.

Gerasimos Potamianos is an Associate Professor in the Department of Electrical and Computer Engineering at the University of Thessaly in Greece and holds a Ph.D. degree from Johns Hopkins University (1994). Prior to his current position, he has been at the Center for Language and Speech Processing at Johns Hopkins, at AT&T Labs-Research, and the IBM T.J. Watson Research Center in the USA, followed by FORTH and Demokritos Research Centers in Greece. His research interests span multisensory and multimodal speech processing and scene analysis with applications to human-computer/-robot interaction and ambient intelligence. He has authored 160 articles that have received over 6.5k citations, holds 7 patents, and has been involved in numerous European and national research projects. He has served as an organizing committee member of EUSIPCO'17, SLT'18, and ICASSP'23, at the IEEE Speech and Language Committee, and is currently a member of IEEE, ISCA, EURASIP, and the Technical Chamber of Greece.

Khasmamad Shabanovi received his B.S. degree in Computer Engineering from Bilkent University. During his senior year, he researched discovering a correlation between pose and apparent personality traits. Currently, he is pursuing an M.S. degree in Computer Science at the Technical University of Munich. His research interests are artificial intelligence and deep learning.

Tal Shafir graduated from law school at the Hebrew University of Jerusalem. Following her passion, she then studied dance-movement therapy at the University of Haifa, and completed her Ph.D. in neurophysiology of motor control and two postdoctoral fellowships: in brain-behavior interactions in infants, and in affective neuroscience, all at the University of Michigan. While working as a research investigator at the University of Michigan, Department of Psychiatry, she started to develop her research on movement-emotion interaction and its underlying brain mechanisms, behavioral expressions, and therapeutic applications, which she now continues at the University of Haifa. Shafir, certified also in Laban Movement Analysis, was the main editor of *The Academic Journal of Creative Arts Therapies*, and of *Frontiers in Psychology* research topic: “The state of the art in creative arts therapies.” She has been serving on The American Dance Therapy Association (ADTA) research committee since 2016, and was the recipient of ADTA 2020 Innovation Award.

Sinan Sonlu received his B.S. and M.S. degrees in Computer Engineering from Bilkent University. He is currently pursuing Ph.D. studies at the same university. His research interests include conversational virtual agents, expressive animation, motion generation, and personality synthesis. With his research group, he currently works on successfully representing the different personality traits in virtual character animation. He and his colleagues recently published their conversational agent framework with multimodal personality expression.

David G. Stork is an Adjunct Professor at Stanford University. He received a B.S. degree in physics from the Massachusetts Institute of Technology, Cambridge, MA, USA in 1976 and a Ph.D. degree in physics from the University of Maryland, College Park, MD, USA in 1984. He has made contributions to machine learning, pattern recognition, computer vision, artificial intelligence, computational optics, image analysis of fine art, and related fields. He is a Fellow of seven international scholarly societies, and his eight books/proceedings volumes include the second edition of *Pattern Classification* and *Pixels and Paintings: Foundations of Computer-Assisted Connoisseurship*.

Natalie Strand is a graduate student at The Pennsylvania State University working with Dr. Reg Adams. She has a B.S. in Behavioral and Cognitive Neuroscience and is currently pursuing a Ph.D. in Psychology with a specialization in cognitive and affective neuroscience. Natalie’s research interests are broadly related to the influence of compound social cues on emotion perception in the face. More specifically, her work bridges social vision and feminist approaches to investigate how structural facial cues and gender/sex emotion stereotypes influence emotion perception within and outside of the gender binary. Natalie is also interested in exploring the use of computer vision models to examine how facial cues impact emotion perception.

Zhangyang Wang is currently the Jack Kilby/Texas Instruments Endowed Assistant Professor of Electrical and Computer Engineering at The University of Texas at Austin. He received his Ph.D. degree in ECE from UIUC in 2016, advised by Professor Thomas S. Huang; and his B.E. degree in EEIS from USTC in 2012. Prof. Wang has broad research interests spanning from the theory to the application aspects of machine learning (ML). At present, his core research mission is to leverage, understand, and expand the role of sparsity, from classical optimization to modern neural networks, whose impacts span many important topics such as efficient training/inference/transfer, robustness and trustworthiness, generative AI, and graph learning. Prof. Wang has received many research awards, including an NSF CAREER Award, an ARO Young Investigator Award, an IEEE AI's 10 To Watch Award, an INNS Aharon Katzir Young Investigator Award, and a few more industry research awards.

James Z. Wang is a Distinguished Professor at The Pennsylvania State University. He received a Bachelor's degree in Mathematics *summa cum laude* from the University of Minnesota (1994) and an M.S. degree in Mathematics (1997), an M.S. degree in Computer Science (1997), and a Ph.D. in Medical Information Sciences (2000), all from Stanford University. His research interests include affective computing, image analysis, image modeling, image retrieval, and their applications. He was a Visiting Professor at the Robotics Institute at Carnegie Mellon University (2007–2008), a lead special section Guest Editor of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2008), and a program manager at the Office of the Director of the National Science Foundation (2011–2012). He is also affiliated with the Department of Communication and Media, School of Social Sciences and Humanities, Loughborough University, UK (2023–2024).

Lai-Kuan Wong is currently an Associate Professor with the Faculty of Computing and Informatics and the Chair of the Center for Visual Computing at the Multimedia University, Malaysia. She received the B.Sc. degree in Computer Science from Universiti Sains Malaysia, Malaysia and the M.Sc. and Ph.D. degrees in Computer Science from the National University of Singapore. Her research interests include computational photography, computational aesthetics, stereo image and video enhancement, and medical imaging. She serves as Co-Chair for several international workshops held in conjunction with Asian Conference on Pattern Recognition 2015, Asian Conference on Computer Vision 2018, and ACM Multimedia 2020, and as the Organizing Committee for several international conferences including International Conference on Image Processing 2023, ACM Multimedia Asia 2023, IEEE International Conference on Multimedia & Expo 2022, International Symposium on Intelligent Signal Processing and Communication Systems 2022, and Workshop on Multimedia Signal Processing 2018.

Benjamin Wortman is a Ph.D. candidate in the Informatics program at The Pennsylvania State University. He received a Bachelor's degree in Data Sciences and an M.S. degree in Informatics from The Pennsylvania State University, University Park, in 2020 and 2022, respectively. His research interests include affective computing, computer vision, and machine learning.

Dejia Xu is a Ph.D. student from the Department of Electrical and Computer Engineering at the University of Texas at Austin, advised by Prof. Zhangyang (Atlas) Wang. He received his B.S. degree from the School of Electronics Engineering and Computer Science at Peking University in 2021. His research interests include computational photography, creative vision, and implicit neural representation. He is one of the recipients of the 2022 Snap Research Fellowship.

Feng Xu received a B.S. degree and an M.S. degree in Computer Science from Fudan University in 2013 and 2016, respectively. His research interests include machine learning, computer vision, and affective computing. He has been engaged in cyber risk management at Ant Financial Group since 2016.

Yifan Yuan received a B.S. degree in Physics from Fudan University in 2019 and is now a Ph.D. student in Computer Science at Fudan University admitted in 2021. Her research interests include image attribute manipulation, microexpression recognition, and image generation.

Junping Zhang received a B.S. degree in Automation from Xiangtan University, China, in 1992, an M.S. degree in Control Theory and Control Engineering from Hunan University, Changsha, China, in 2000, and a Ph.D. degree in Intelligent Systems and Pattern Recognition from the Institution of Automation, Chinese Academy of Sciences, in 2003. He has been a Professor at the School of Computer Science, Fudan University, since 2006. His research interests include machine learning, image processing, biometric authentication, and intelligent transportation systems. He has been an Associate Editor of *IEEE Intelligent Systems* since 2009 and was an Associate Editor of *IEEE Transactions on Intelligent Transportation Systems* (2010–2018).

Sitao Zhang is a Ph.D. candidate in the Informatics program at The Pennsylvania State University, advised by James Z. Wang. His primary focus of research centers on the fields of computer vision and machine learning, with a specific emphasis on vision-language integration and self-supervised learning. Before joining Penn State, he received a Master's degree in Mathematics from the University of Wisconsin-Madison and a Bachelor's degree in Statistics from Sun Yat-sen University.

Acronyms

AAM	Active Appearance Model
ACG	Attributed Composition Graph
ACLU	American Civil Liberties Union
ACM	Association for Computing Machinery
ADTA	American Dance Therapy Association
AEI	Artificial Emotional Intelligence
AI	Artificial Intelligence
AMT	Amazon Mechanical Turk
ANE	Apple Neural Engine
ANOVA	Analysis of Variance
APA	American Psychological Association
ARO	Army Research Office
ASA	American Statistical Association
ASD	Autism Spectrum Disorders
ASM	Active Shape Model
AU	Action Unit
AUC	Area Under the Curve
AUC ROC	Area Under the Receiver Operating Characteristic Curve
BEEU	Bodily Expressed Emotion Understanding
BRISQUE	Blind/Referenceless Image Spatial Quality Evaluator
BoLD	Body Language Dataset
CMA	Certified Movement Analyst
CNN	Convolutional Neural Network
COCO	Common Objects in Context
CRI	Child-Robot Interaction
CT-MC	Continuous Time Markov Chain
DARPA	Defense Advanced Research Projects Agency
DCT	Discrete Cosine Transform
DNN	Deep Neural Network
DSLR	Digital Single-Lens Reflex
DT-MC	Discrete Time Markov Chain

EDR	Endpoint Detection and Response
EEG	Electroencephalography
EMFACS	Emotion Facial Action Coding System
FACS	Facial Action Coding System
fMRI	Functional Magnetic Resonance Imaging
FPS	Frame-per-Second
GAN	Generative Adversarial Network
GPT	Generative Pre-trained Transformer
GPU	Graphics Processing Unit
HDR	High Dynamic Range
HEIC	High Efficiency Image Container
HICEM	High-Coverage Emotion Model
HRI	Human-Robot Interaction
HSI	Hue, Saturation, Intensity
I/O	Input and Output
IAA	Image Aesthetics Assessment
IQA	Image Quality Assessment
IEEE	Institute of Electrical and Electronics Engineers
ISP	Image Signal Processor
JPEG/JPG	Joint Photographic Experts Group
LBMS	Laban/Bartenieff Movement System
LDA	Latent Dirichlet Allocation
LMA	Laban Movement Analysis
MEG	Magnetoencephalography
ML	Machine Learning
MP	McCulloch-Pitts
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NIQE	Natural Image Quality Evaluator
NLP	Natural Language Processing
NSF	National Science Foundation
OCEAN	Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism
OEM	Original Equipment Manufacturer
PAD	Pleasure, Arousal, and Dominance
PCA	Principal Component Analysis
PLD	Point-Light Display
PSNR	Peak Signal-to-Noise Ratio
RAM	Random-Access Memory
RBF	Radial Basis Function
RGB	Red, Green, and Blue
RYB	Red, Yellow, and Blue
SES	Socio-Economic Status
SVM	Support Vector Machine
SVR	Support Vector Regression

TIPI	Ten Item Personality Inventory
TSN	Temporal Segment Network
UI/UX	User Interface and User Experience
VAD	Valence, Arousal, and Dominance
VR	Virtual Reality

Part I
Foundations of Emotion Modeling and Machine Learning

Because this book is multidisciplinary in nature, this part will provide essential knowledge on emotion models and machine learning fundamentals.

Chapter 1, "Models of Human Emotion and Artificial Emotional Intelligence," aims to bridge the gap between emotion models used in psychology and their application in affective computing tasks. It surveys existing emotion models in psychology, highlighting their strengths and weaknesses for computational tasks involving human emotion.

Chapter 2, "A Concise Introduction to Machine Learning," offers a fundamental understanding of machine learning techniques relevant to the theme of the book. This chapter serves as a starting point for readers with limited or no relevant expertise. It introduces learning algorithms, basic concepts, and fundamental principles in machine learning systems.

These chapters serve as a valuable foundation for readers interested in emotion modeling and the application of machine learning in aesthetics, emotion, and artistic style. By familiarizing themselves with emotion models and understanding the basics of machine learning, readers can better comprehend and engage with recent research in these areas presented in the rest of the book and beyond.

Chapter 1
Models of Human Emotion and Artificial Emotional Intelligence

Benjamin Wortman

Abstract This chapter bridges the gap between emotion models popular in psychology and their use in affective computing tasks. Emotion modeling has a long and varied history with several competing schools of thought. Here, through a survey of existing literature, we cover existing emotion models popular in psychology, highlighting the strengths and weaknesses of these different approaches in regard to computational tasks involving human emotion.

1.1 Introduction

As an interdisciplinary field, affective computing sits at the corner of psychology and computing. This field has seen exponential growth in recent years[1] in tandem with the rise of Deep Learning and the creation of large scale datasets for training [10, 22, 33, 40, 51]. Although historically the focus has been on developing algorithms and techniques to help machines recognize and respond to human emotion, these technologies are underpinned by models in psychology. This is important to understand when transitioning to real-world settings as ultimately a machine-learning model is limited by the usefulness of the underlying emotion model. For example, Ekman's universal basic emotions [24] has long been the dominant model used for comparison in this field. In a lab setting, this has several benefits such as its ease of use and cross-cultural relevance. However, given it only consists of 7 emotions, this narrow coverage makes it difficult to develop real-world applications with high enough fidelity to accurately describe the wide variety of emotional states that humans present.

[1] Over 14,000 emotion recognition papers since 2010 according to IEEE Xplore.

B. Wortman (✉)
College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, USA

J. Z. Wang, R. B. Adams, Jr. (eds.), *Modeling Visual Aesthetics, Emotion, and Artistic Style*, https://doi.org/10.1007/978-3-031-50269-9_1